Topsy Scoring System

Understanding How Topsy Checks for Relevancy and the Application of Topsy as a Scholastic Grading Tool

Karl Rozits
Adam Brown
CEE 300
Dr. Thomas Seager
April 23, 2013

Executive Summary

Arizona State University's CEE300 course on business practices is distinguishable from the other engineering courses offered at ASU for many reasons. One of the most distinguishable part of the class is the use of Twitter as a means of gauging class participation, which makes up 100 of the total 900 points of the students overall letter grade. Students may also earn experience points (XP), which is a form of extra credit. In an attempt to measure the relevancy of students' tweets to their relation to the course objectives, an online tool known as Topsy is used. Topsy is an online analytical tool that evaluates millions of archived and real-time tweets based on their relevancy to a specific criterion. For gauging a student's participation in CEE300, the students' unique Twitter handle is simply entered into Topsy's search engine along with the hashtag #CEE300. The results are populated within seconds and can be analyzed according to many factors, including date, time, relevancy, tweets, photos, videos, etc. The results are then directly used as part of the students' class participation grade. The first 75 tweets made count towards the participation grade, while any number of tweets above 75 are counted as XP points. The grading system is modeled so that XP are subject to diminishing returns, meaning the more XP that is earned, the less an individual experience point is worth. In addition, XP will help a student who has a poor grade more than someone who has a high grade due to the exponential behavior of the grading model. The objective of the report is to determine what Topsy considers relevant, to understand how to create a relevant tweet, to verify the accuracy of Topsy's relevancy score and to conclude whether Topsy is an acceptable tool for use in gauging class participation. After thorough investigation, Topsy was determined to be a great analytical tool for monitoring Twitter participation, yet lacks the fundamental ability to distinguish between tweets relevant to CEE300 and tweets relevant to everything else. The report recommends that a policy change of the value

of each Topsy score be weighted as a fraction of an experience point, thus minimizing the desire

of students to make an attempt to increase their Topsy score. Doing so also encourages other

forms of class participation such as commenting on the *Sustainable Engineering Systems* blog

posts by Dr. Seager.

Table of Contents

List of Tables

List of Figures

Introduction

Twitter has become a large part of American culture in recent years. From personal use as a social media outlet to business use for attracting consumers, Twitter has a wide range of applications. For the purposes of CEE300, Twitter is used for three reasons.

The first reason is to encourage class participation. For many students, the thought of participating in class is not appealing; either they do not care to be active, do not want to share their ideas, or they are too shy to speak in front of others. Twitter allows these students to be active with course content and socialize with classmates. Although online participation does not take the place of in-person participation, Twitter keeps students attentive to course material.

The second reason is that the grading of tweets can be similar to that of a popularity contest. As stated in the *Executive Summary* section, Topsy tracks relevant tweets. The criteria for relevancy will be discussed later in the report, but for now relevancy and popularity mean the same thing. Although popularity may not be the first thought when thinking of a professional work environment, how much a person is liked will affect their progress in their career. Because Topsy tracks popular tweets, the students who are more popular will benefit more from their tweets than students who Topsy does not deem popular. Given that grades supposed to reflect academic and not social success, many students thought the grading system for Topsy was unfair.

The third reason is social media such as Twitter is becoming a more integrated part of the business world. Social media allows people to connect with each other. Since the workplace is a time for work, Twitter allows coworkers to get to know each other more outside of work as well.

Students do not need to worry about the issue since they have plenty of opportunities to socialize at school, but getting used to social media could make their future connections easier to establish.

The following report will analyze the application of Twitter and Topsy as a means to grade students for participation. At the end of the report, a recommendation to the instructor of CEE300, Dr. Thomas Seager, will be made with regards to what should be changed about the use of Twitter and Topsy in the class.

What is Relevancy?

According to the Merriam-Webster dictionary, relevancy is defined as having significant and demonstrable bearing on the matter at hand. In CEE300, relevant tweets pertain to the content being discussed at the time of the tweet. For example, during the module on net present value, tweeting or retweeting about the TEG game played earlier in the semester would not be considered relevant because that is not the focus of class at the time. Tweeting about an earlier module about cash flow diagrams could be considered relevant during the net present value module because cash flow diagrams are related to net present value.

For Topsy, relevancy has to do with the content of a tweet and who posts the tweet. As far as content goes, Topsy tracks website links that are included in tweets. According to Topsy's Social Analytics page, Topsy measures the relevancy of links with three criteria: influence, velocity, and momentum. Influence measures the likelihood that someone is paying attention to, or tweeting about, a link. Velocity measures the speed of increasing popularity of a link.

Momentum measures the popularity and velocity of a link, i.e. the rate at which a link is becoming more popular.

Ideally, students would create tweets that are relevant to CEE300 and not worry about Topsy. However, due to the fact that Topsy is used to find their online participation grade, students are enticed to make their tweets relevant to Topsy. If a student wants to make their tweets relevant to Topsy, they must include popular links in their tweets. Although that method may work, Topsy also takes into consideration the popularity of the person tweeting.

Establishing Relevancy

As mentioned in the previous section, Topsy measures the relevancy of not only links, but also people. Just like in the real world, people pay attention more to those who are popular than those who are not. Most students in CEE300 had to make a Twitter account for the class because they did not have one previously. Because most students had new accounts at the beginning of the semester, they were not considered relevant by Topsy and were therefore less likely to receive hits on Topsy. Towards the end of the semester, the majority of the class was able to make themselves more relevant to Topsy. How were they able to make their tweets more relevant? Students had to establish a level of relevancy with Topsy in order to receive more hits from Topsy.

Establishing relevancy can be accomplished by two methods: the expedited and consistent methods. The expedited method is to make tweets filled with relevant links. If Topsy sees that a person is talking about things that are relevant, Topsy will attribute a level of

relevancy to that person. Using the expedited method will allow a person to be relevant quickly. The long-term method is to be consistently active and making relevant tweets over a spread out period of time. Having consistent and regular tweeting habits will show Topsy that a person is relevant, as long as other people are interacting with them and they make tweets that have some sort of relevancy to Topsy. *Figure 1*, below, shows a comparison between the expedited and consistent methods.

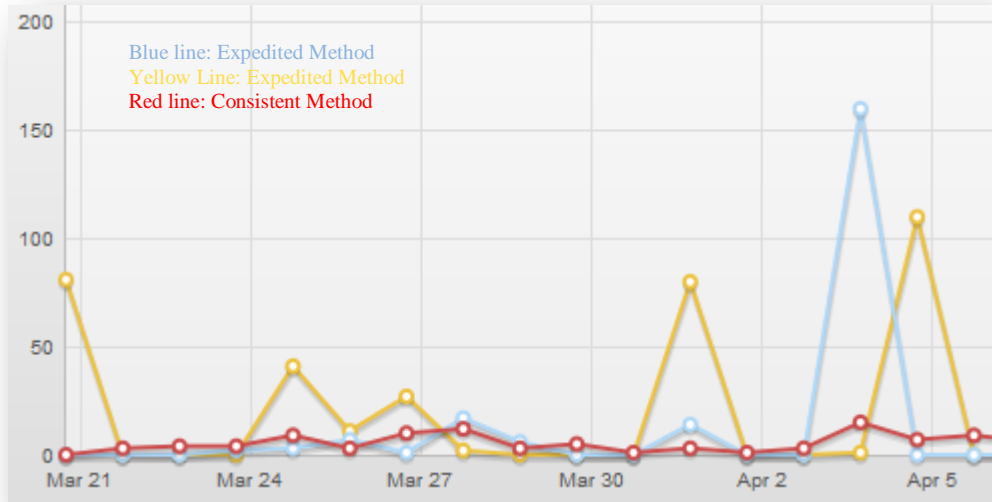*Figure 1: Comparison of the Expedited and Consistent Methods*

The blue and yellow lines represent two different Twitter accounts that use the expedited method. The red line represents a Twitter account that uses the consistent method. *Figure 1* shows that both methods result in Topsy hits, the only difference is that the expedited method results in relevancy faster than the consistent method.

4

Investigative Method

The primary tool for the analysis of Topsy scores was undoubtedly the Topsy Social Search website, along with two valid Twitter accounts. Throughout the study, numerous tweets were submitted via Twitter and the results of these tweets were analyzed on Topsy under the "Everything" search option, sorted by "Relevance", on the "Twitter" network and in "All Languages". Some of the tweets submitted were meaningless and unrelated to CEE300; others were related to CEE300 and meaningful, while still others met a handful of different conditions. The tweeting trials used in determining the variety of tweets to submit are listed in *Table 1* and are described in the *Results* section of the report. Once the tweets were made, the results obtained from Topsy were analyzed to find a correlation between the trials' characteristics and applicable relevancy score. The relationship was then used to recommend a more efficient participation grading system for CEE300.

Table 1: Trial Descriptions

Trial	Relates to CEE300	Unrelated to CEE300				Includes a Link	Retweet
		Random Tweet	Topsy top 100	Topsy SQ	Twitter Trend		
CEE300, Link	X					X	
CEE300, No Link	X						
Random, Link		X				X	
Random, No Link		X					
Top 100, Link			X			X	
CEE300, Link, Retweet	X					X	X
Random, Link, Retweet		X				X	X
Top 100, Link, Retweet			X			X	X
Twitter Trend					X		
CEE300, Random, Retweet	X	X					X
Topsy Suggested Queries				X			

Results

As listed in *Table 1*, there were many different tweeting conditions that were tested throughout the experiment in order to further understand how Topsy estimates relevancy. The trial abbreviations from *Table 2* are used when describing the results of the investigation. For trials CL and CnL 20 tweets were posted through a CEE300 student account via Twitter, all of which related to the course subject matter of CEE300. 10 of those tweets included a link to the blog videos associated with the course, while the other 10 did not. Of the 10 that did include a link of the blog video, 9 of the posts were deemed relevant by Topsy, compared to only 2 without a link, suggesting that a more relevant tweet is one that includes a link.

Table 2: Trial Results

Trial	Trial Abbreviations	Number of Tweets	Number of Topsy Hits	Accuracy
CEE300, Link	CL	10	9	90%
CEE300, No Link	CnL	10	2	20%
Random, Link	RL	20	16	80%
Random, No Link	RnL	20	1	5%
Top 100, Link	100L	20	17	85%
CEE300, Link, Retweet	CLRt	10	10	100%
Random, Link, Retweet	RLRt	10	10	100%
Top 100, Link, Retweet	100LRt	20	20	100%
Twitter Trend	TwT	10	1	10%
CEE300, Random, Retweet	CRRt	10	7	70%
Topsy Suggested Queries	ToSQ	10	2	20%

TOPSY SCORING SYSTEM

For trials RL, RnL and 100L, 60 tweets were posted in a similar manner, yet none were related to the course material. Of these 60 tweets, 40 were randomized tweets without any significance to anything, while the other 20 contained links to websites taken directly from the *Trending 100 Links* page on Topsy.com. Of the 40 random tweets, 20 contained links to random non-trending websites while the other 20 did not contain any link at all. Only 17 of the trending tweets were deemed relevant by Topsy. Comparatively, 16 of the non-trending links were deemed relevant by Topsy as opposed to only 1 that did not include a link.

Trials CLRt, RLRt, 100LRt and CRRt focused on the retweet function of Twitter. Retweeting is simply a way for people to forward someone else's tweet and attaching the retweeter's Twitter handle to the tweet. For CLRt trial, 10 tweets that were posted from other CEE300 students containing links to the course blog were retweeted. Of these, all 10 showed up on Topsy. For the RLRt trial, 10 tweets were retweeted from other students that had no relation to the course material. Of these, all 10 were counted by Topsy. For the 100LRt trial, 10 tweets were retweeted from another student that had tweeted trending topics, yet had no relation to CEE300. Once again, all 10 were deemed relevant on Topsy. Lastly, for the CRRt trial, 10 tweets about the class blog videos were retweeted from another student; the 10 tweets did not show up in the students' Topsy search results. Of these, 7 of the retweets showed up on Topsy.

Trials TwT and ToSQ focused on Topsy suggested queries and Twitter trends, respectively. Topsy suggested queries are similar to Google suggested queries; when characters are typed into the search bar on Topsy, the search engine suggest the most popular search query that starts with the character. For the TwT trial, 10 tweets were made with the suggested query options and only 2 tweet showed up on Topsy, both of which had been retweeted by others. Twitter trends are found on the home page of Twitter and can be altered to show suggested

trends for the user or to show worldwide trends. For the ToSQ trial, 8 tweets were made with the user suggested trends and 2 tweets were made with worldwide trends, totaling 10 tweets. Of the 10 tweets made, only 1 showed up in the Topsy search results; the tweet was retweeted by someone else.

  Figure 2, below, shows the individual results of the trials graphed in a radar plot. A radar plot was used for easy contrast between the comparable, yet different trials. The data used to make the graph can be found in *Table 2* under the "Accuracy" column.
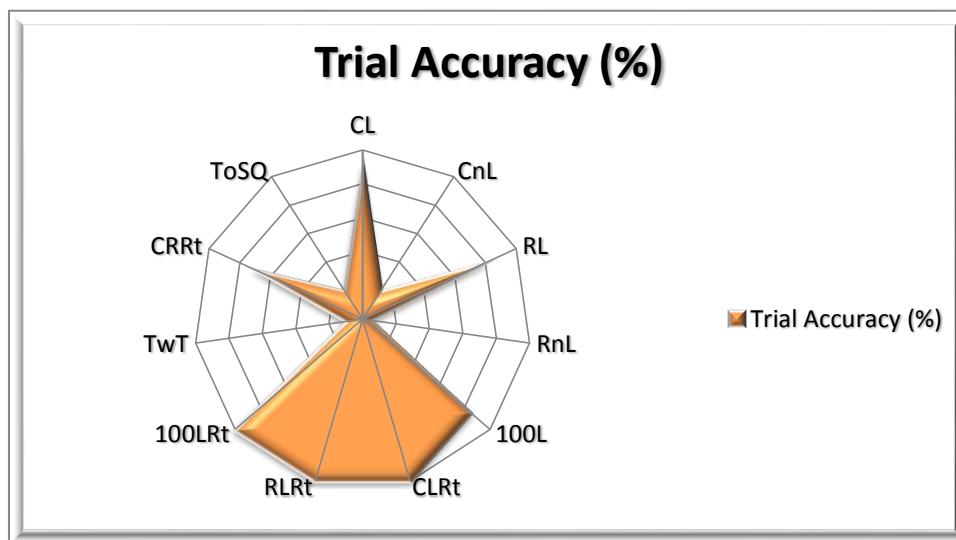

Figure 2: Radar plot of Trial Accuracy

  Another piece to the Topsy puzzle is how the popularity of the tweeter affects the number of hits from Topsy. For the experiment, an additional Twitter accounts (Account A) was created to test a few hypotheses. Account A was created and immediately posted 5 non-trending tweets; all of which did not register on Topsy. Account A then tweeted 5 trending topics and contrary to

the above information, there were still no signs of relevancy on Topsy. Once an established and active student Twitter user started retweeting Account A's tweets, Topsy deemed the account relevant and began counting the trending tweets made. Account A then sat idle for a day before continuing to tweet, at which point Topsy decay took place and dropped Account A's score by an arbitrary amount. Interestingly, the Topsy score from the established Twitter user did not experience the decay for the tweets that were retweeted from Account A. Topsy decided that Account A was no longer popular, thus the future tweets from Account A were not displayed on Topsy. Once a separate, established account began retweeting Account A's tweets, Account A regained relevancy on Topsy.

An interesting phenomenon occurs when someone who has established relevancy retweets a tweet made by someone who has not established relevancy (trial CRRt). The relevant tweeter will have the retweet show up in their Topsy search results, but the non-relevant tweeter may not always receive the same benefit. In fact, even if a tweet that previously did not show up on Topsy is retweeted by a relevant person, that tweet will still show up on the relevant person's Topsy search results. The result is actually supported by how the real world works; people are more likely to pay attention to others if they are popular and less likely to pay attention to others if they are not popular.

Another point to be made from the results is about Twitter user options such as "Reply" and "Favorite." When making the test tweets, the testers had no control over the people following them from using these options, along with the "Retweet" option. Of the 150 total test tweets made, 9 retweets, 5 replies and 3 favorites were received on a mixture of tweets. All 9 tweets that received retweets displayed in Topsy search results. None of the 5 replies and 3 favorites showed up in the search results. From the results, we can conclude that retweeting

boosts the relevancy of a tweet, while favoriting and replying has no effect on the relevancy of a tweet whatsoever.

Tweeting and the Power Law (80-20 Rule)

The power law (a.k.a. the 80-20 Rule) is a probability distribution in which the top 20% of the population observed is responsible for approximately 80% of the data recorded. The power law is based off the function $f(x) = c^x$ where x represents the number of tweets made and c represents the scaling factor to make the area under the curve equal to one. The function was not used, as the data points were already available. The larger the population being observed, the more accurate the power law distribution becomes (up to a certain point). With the small population size of CEE300, 68 students, the power law is not as accurate, but still acts in the same fashion as a larger population. As seen in *Figure 3*, the top 20% of students (14 students) that tweeted made 4,862 tweets of the total 10,142 tweets (see *Appendix, Table 3* for data).
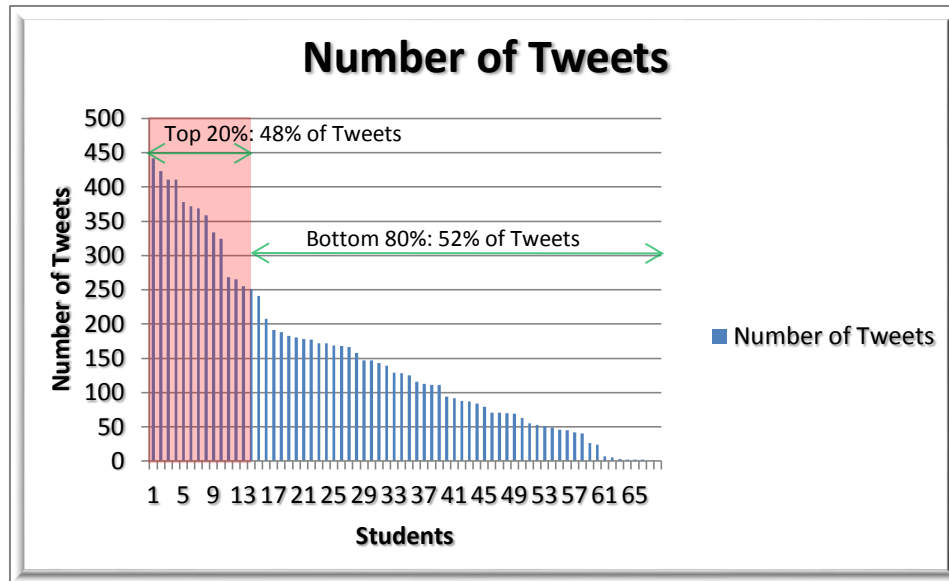
Figure 3: Power Law: The number of tweets made in the spring 2013 CEE300 class

Figure 3 shows that 48% of the number of tweets was made by the top 20% of students in class while the remaining 52% of tweets were made by the bottom 80% of students. In the case of CEE300, the power law means there are a few students (20%) who make most of the tweets in class, while the majority of the class (80%) makes up the remainder. In other words, for an open-ended assignment such as tweeting, there are only a few students who are driven to do their best or to obtain the most amount points. The average number of tweets made in class was 149 tweets per person. Only 28 students were shown to be above average, while the remaining majority of 40 students are below average. The average is a misleading number though, as the average is taken from two extremes of high and low volumes of tweeting. If an instructor grades tweeting on a curve, the majority of students who are "below average" will not benefit because the high tweeting students set the curve so high that the lower tweeting students will not see an increase in their grades. Proof that the power law is applicable is seen in *Figure 4* below.
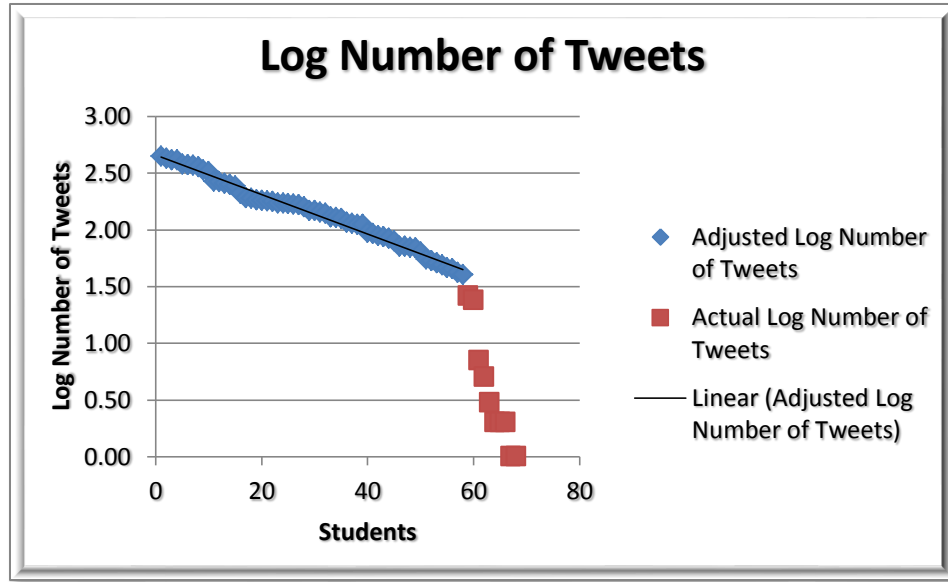
Figure 4: Power Law: Log of the number of tweets made in the spring 2013 CEE300 class

Figure 4 shows the linear relationship between students and the logarithmic values of the number of tweets they made. The "Actual Log Number of Tweets" series was separated from the "Adjusted Log Number of Tweets" to make a better linear fit for the data. In a larger population, these few data points would not be as apparent because of the comparatively larger number of data points. Since the CEE300 class is a small population, the "Actual Log Number of Tweets" data points were more obvious. Although the application of the power rule does not follow the probability distribution verbatim, the power rule is still pertinent to the analysis.

Making Relevant Tweets for Topsy

As discussed in the *What is Relevancy?* section, Topsy has a certain criteria for determining relevancy. The criterion includes the level of influence, momentum and velocity of

13

the contents of a tweet and the person making the tweet. *Figure 4* shows a proposed algorithm

for how Topsy determines relevancy. The "R%" represents the minimal percentage of relevancy

a tweet must contain in order to display in the search results on Topsy. Note that user options on

Twitter, such as retweeting, are not a part of the algorithm. Retweeting plays a part in all three

aspects in the algorithm (influence, momentum and velocity) and therefore is not a separable
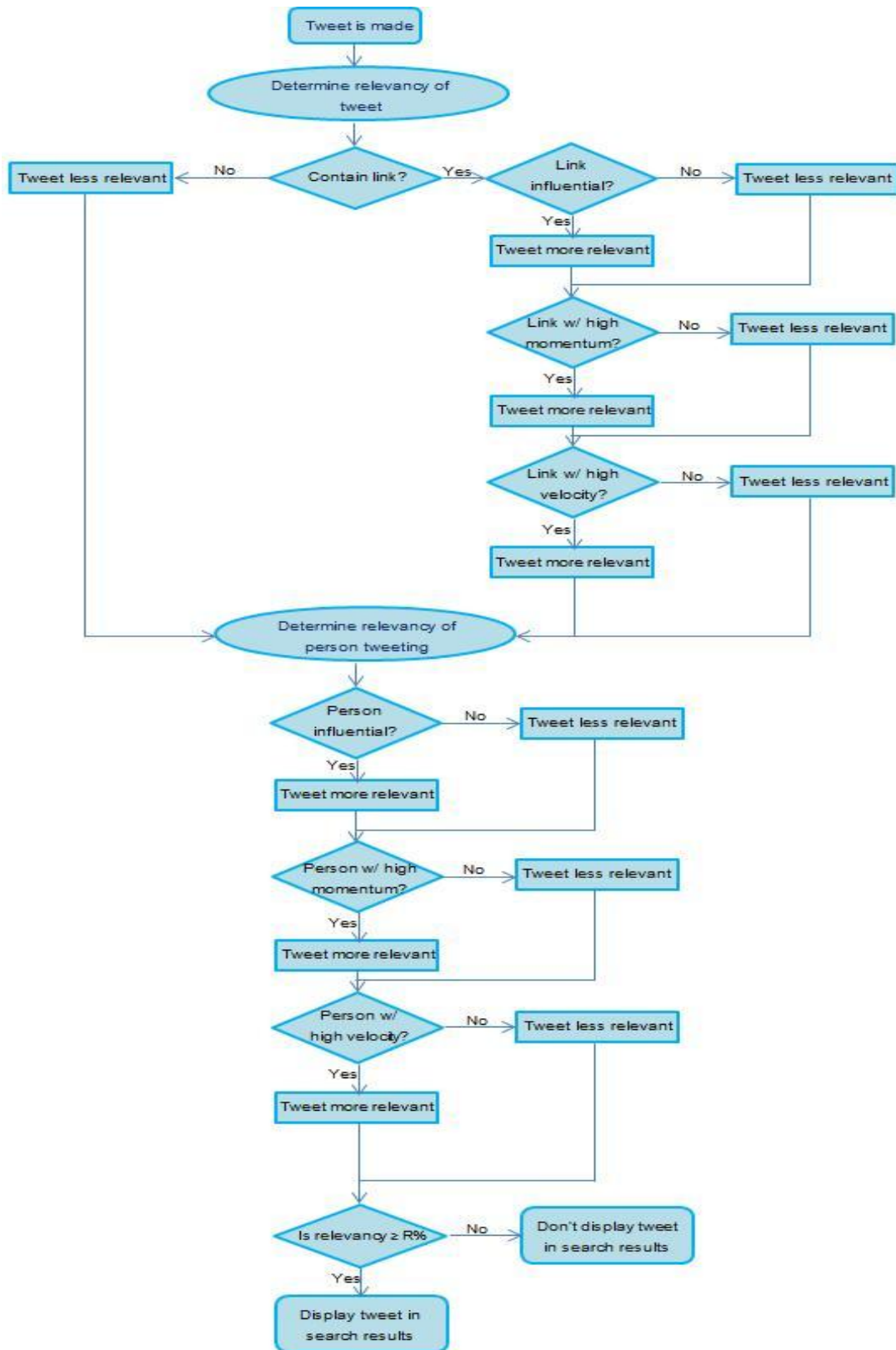
check to make.

Figure 5: Proposed Topsy Algorithm

Following the above proposed algorithm can lead to success when trying to make tweets relevant to Topsy. Although the actual Topsy algorithm is likely to be more detailed and complex, the proposed algorithm is simple enough where anyone can have a basic understanding of how Topsy determines relevancy.

Below, *Figure 4* shows a tweet that has been made relevant to Topsy.

Figure 6: A tweet relevant to Topsy

The tweet includes a link from the *Trending 100 Links* page on Topsy.com. A trending link means that the link fulfills all or some of the categories that Topsy checks for relevancy. The tweet in *Figure 4* did show up in the Topsy search results and the student may therefore receive credit for the tweet. *Figure 5* shows a tweet that has not been made relevant to Topsy.
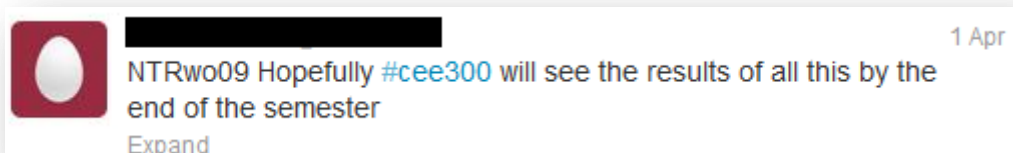

Figure 7: A tweet not relevant to Topsy

The tweet contains no links or additional relevancy aside from the relevancy of the person who made the tweet. Topsy has not calculated the person who made the tweet relevant enough for the tweet in *Figure 5* to show up in Topsy results so the student will not receive credit for the tweet.

Making Relevant Tweets for CEE300

Relevancy for CEE300 is described in the *What is Relevancy?* section earlier in the report. Making tweets about course content that is being discussed at the time along with making tweets connecting current and previous course content are tweets that are relevant to CEE300. *Figure 6*, below, shows a tweet that has been made relevant to CEE300.
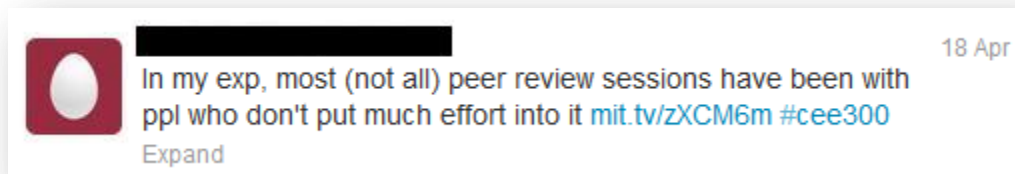

Figure 8: A tweet made relevant to CEE300

The tweet in *Figure 6* is addressing the course content at the time while including a link to a blog video and the CEE300 hashtag. Ideally, tweets that fit the format above are what students should be making. *Figure 7*, however, shows a tweet that is not relevant to CEE300.

Figure 9: A tweet not made relevant to CEE300

The tweet in *Figure 7* shows a tweet that has not been made relevant to CEE300, but still includes the CEE300 hashtag and a link. Clearly, the student should not be given credit for the tweet as part of their online participation grade; if the tweet shows up in Topsy, however, the student could very well receive the credit. The tweet above has been formatted to appeal to Topsy, but not to CEE300.

Making Relevant Tweets for Both Topsy and CEE300

As seen in the previous two sections, students can make their tweets appeal to Topsy or CEE300 very easily. The challenge students face is finding the correct way to make their tweets relevant to both. A solution is not difficult to find though, as seen in *Figure 8.*
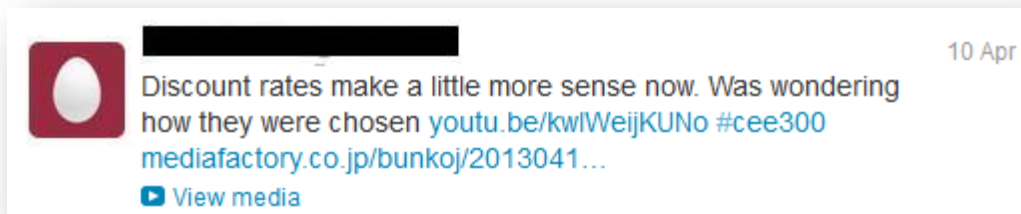

Figure 10: A tweet made relevant to both Topsy and CEE300

The tweet in *Figure 8* shows a tweet that has been made relevant towards both Topsy and CEE300. The tweet contains conversation about current course content with a link to the blog video, the CEE300 hashtag and an additional link that was found from the *Trending 100 Links* page on Topsy.com. Given that the tweet contains content in CEE300, the tweet can be counted fairly by the instructor. If students choose to make their tweets relevant to both Topsy and CEE300, they will be talking about course content more and improving their grades; the outcome is beneficial to the students as they learn more and receive a better grade in the course.

Topsy Decay

Topsy decay is the term that describes occurrences when a tweet displays in Topsy search results shortly after the tweet is made, then does not display in the Topsy search results after a longer amount of time. Students noticed that Topsy determines the relevancy of tweets within a certain amount of time. Most tweets made will display on Topsy very quickly after they are made, but if there are a large amount of tweets being made for the search query, not all may show right away. If a tweet still displayed in Topsy search results after 24 hours, there was a decent chance that tweet would remain there. If a tweet does not display after 24 hours, Topsy determined that the tweet was not relevant. Topsy decay can be affected by the amount of relevancy of both the tweet and the person making the tweet. For most of the students of CEE300, a Twitter account had to be made for the class, meaning the students had no previous relevancy to Topsy. On top of that, the students were unaware of how Topsy determined relevancy so their tweets were less likely to show up in Topsy search results. After Topsy decay had been discovered, many students began searching for ways to make their tweets relevant to Topsy. As the students found methods for doing so, almost all students in CEE300 started using the methods to boost their Topsy hits and thus reducing Topsy decay. Although Topsy decay was not a worry to the students of CEE300 by the end of the semester, other problems arose due to the methods being used to increase students' Topsy hits.

Problems with Student Knowledge of Topsy Relevancy

The data above clearly shows that Topsy decay can be easily combatted by making tweets more relevant to Topsy. Likewise, retweeting any sort of tweet is likely to score a hit on

Topsy. Consequently, the simplest method of obtaining and maintaining a relevant score on Topsy is to retweet a previous tweet, regardless of the content. Retweeting has an obvious downfall; a student may simply retweet tweets relevant only to Topsy back-to-back with no relationship to CEE300 course material and will receive participation points for them. A student posting tweets relevant to the course material may only receive a fraction of the participation points of the retweeting student due to the lower amount of Topsy hits they receive.

The outcome of retweeting is likely to motivate other students to neglect the actual course material and retweet tweets relevant only to Topsy as well, leading to an overabundance of spam-like tweets on the Twitter feeds of all of the students. Student abuse of retweeting was experienced in CEE300 near the end of the semester and was discussed in class. Many students, the ones who were responsible for the mass amounts of retweets, did not see any problem with retweeting, as doing so boosted their grades. For others who used Twitter for personal use and to genuinely be active in class were bothered by the large amount of retweets as they forced the users to sift through their Twitter feed. As of the completion of the report, no outcome has been made from the conversation in class about retweeting.

Aside from negatively affecting Twitter users, sifting through the hundreds of tweets on Topsy to find tweets relevant to the class would be a time consuming experience for the class graders. While Topsy reports the number of relevant tweets a person makes, Topsy does not display every individual tweet they report. As seen in *Figure 9*, Topsy may report that a person has 92 relevant tweets, but when going through the detailed search results, only 62 tweets will show.

Figure 11: Topsy search results error

If Topsy does not show each individual tweet that is reported, then if graders go through all the search results to find the class relevant tweets people make, the students may end up with a lower tweet count since the graders will not be able to see all the relevant tweets made. Another downside to looking through Topsy search results for class-relevant tweets is that many students only retweeting tweets that were relevant to CEE300 earlier in the semester. Although Topsy results clearly distinguish if someone posted an original tweet or simply retweeted something (by showing the retweet under the Twitter account of the original tweeter), if graders chose to give credit for acceptable retweets, they would need to match the date of the retweet with the date of when the course content was being taught.

Final Recommendation

Topsy is no doubt a great analytical tool for evaluating many variables of social networking. However, Topsy's inability to distinguish between spam-like tweets and tweets that are relevant to CEE300 is a significant downfall when gauging class participation. Nevertheless, the students of CEE300 have become familiar with the benefits of using social media such as Twitter, which is a skill they will take with them to the professional world after graduation. Furthermore, although the students' tweets are sometimes irrelevant to the course material, tweeting still motivates interactions between classmates and gives an equal voice to an otherwise quiet student. Even though Topsy may not be the most efficient means of gauging class participation, Twitter is still a beneficial tool for the students.

From an instructor's perspective, counting tweets that are solely attractive to Topsy towards class participation points is less than ideal. The tweets do not show that the student is engaged in class or the material the class is focused on. Students making tweets that are only attractive to Topsy show only that students can inflate the number of tweets they have and therefore collect more XP points at the end of the semester.

After analyzing the results, a number of recommendations could be made. Inserting a clause in the course syllabus stating "no tweet spamming," could stop the problem of mass retweeting completely. The downside to including the clause would be that students would miss out on an opportunity to problem-solve and have valuable dialogue. Going another direction, have the graders enter keywords relevant to course material when looking up the number of Topsy hits of students. Again, there are downsides: doing so will increase the amount of work for the graders and may not give credit to students who do not have the specific keywords in their tweets. Dr. Seager could even discount the amount of Topsy hits for all students based

upon an arbitrary number that depends on how much retweeting happens in the class. Students would argue with the method, though, as they would be likely to lose more points than if they were to not retweet at all; since students cannot control how many tweets other students make, using the method could result in a witch-hunt if there are a few point greedy students. However, a benefit to the method would be to see how students react and compare their reactions to how they act in The Externalities Game played early on in the semester.

Although the above recommendations could work, some better than others, they are still not recommended for implementation. Instead, Dr. Seager should decrease the value of XP earned from tweets in order to discourage students from tweeting or retweeting mass amounts of spam-like tweets. Students can earn XP from other venues, such as commenting on the *Sustainable Engineering Systems* blog, doing practice homework problems on OpenStaxTutor.org, participating in events and research studies or making a contract with Dr. Seager to complete a project he wants done. Because there are plenty of other ways for students to earn XP, and tweeting or retweeting spam-like tweets can become a problem when trying to reward students for legitimate class participation, lowering the amount of XP received from tweeting is a reasonable solution. How much should the XP amount be lowered to? Anywhere from one-tenth to one-fourth of an experience point would be likely to prevent students from tweeting or retweeting mass amounts of spam-like tweets.

Works Cited

Merriam-Webster. (2013). *Relevant - Definition and More*. Retrieved April 20, 2013, from The Free Merriam-Webster Dictionary: http://www.merriam-webster.com/dictionary/relevant

Merriam-Webster. (2013). *Time - Definition and More*. Retrieved April 20, 2013, from The Free Merriam-Webster Dictionary: http://www.merriam-webster.com/dictionary/time

TED Conferences, L. (2005, July). Clay Shirky: Institutions vs. collaboration.

Topsy Labs, Inc. (2012). *Topsy >> Free Analytics*. Retrieved April 20, 2013, from Topsy: http://analytics.topsy.com/

Topsy Labs, Inc. (2012). *Trending 100 Links*. Retrieved April 21, 2013, from Topsy: http://topsy.com/top100

Topsy Labs, Inc. (2013). *Topsy - Instant social insight*. (T. Labs, Producer) Retrieved April 18, 2013, from Topsy: www.topsy.com

Twitter. (2013). *Twitter*. Retrieved April 18, 2013, from Twitter: www.twitter.com

Appendix

Table 3: Power Law Data: Tweets and Log Tweets as of April 19, 2013

Tweets 19 Apr	Log Tweets 19 Apr
442	2.64542
423	2.62634
411	2.61384
411	2.61384
378	2.57749
372	2.57054
369	2.56703
359	2.55509
334	2.52375
324	2.51055
268	2.42813
265	2.42325
255	2.40654
251	2.39967
241	2.38202
208	2.31806
191	2.28103
188	2.27416
183	2.26245
180	2.25527
178	2.25042
177	2.24797
172	2.23553
172	2.23553
169	2.22789
168	2.22531
166	2.22011
158	2.19866
147	2.16732
147	2.16732
143	2.15534
139	2.14301
129	2.11059
128	2.10721
125	2.09691

TOPSY SCORING SYSTEM

116	2.06446
113	2.05308
111	2.04532
111	2.04532
94	1.97313
92	1.96379
88	1.94448
87	1.93952
84	1.92428
79	1.89763
71	1.85126
71	1.85126
70	1.84510
69	1.83885
63	1.79934
55	1.74036
53	1.72428
51	1.70757
49	1.69020
46	1.66276
45	1.65321
42	1.62325
40	1.60206
26	1.41497
24	1.38021
7	0.84510
5	0.69897
3	0.47712
2	0.30103
2	0.30103
2	0.30103
0	-
0	-